

Jaden Wilkins, jwilkins@bellarmine.edu

Jonathan Penaloza Rumie, jpenaloza@bellarmine.edu

- I. **Introduction:** This dataset highlights all the songs that charted in the top 200 between 2020 and 2021. We found the dataset on Kaggle and took an interest in it because of Jonathan's music background.
- II. **Data Set Description:** The dataset originally contained 23 columns, so we added two more. The first column we added was titled 'joyful' which was calculated by adding the provided 'Valence' and 'Danceability' columns. The second column we added was called 'popularity' and was calculated by subtracting 'Artist Followers' from 'Streams'. This was done to highlight popular songs from lesser-known artists.

Variable Name	Data Type	Missing Data %
Highest Charting Position	Interval, int64	0
Number of Times Charted	Int64	0
Week of Highest Charting	Int64	0
Song Name	Object	0
Streams	Int64	0
Artist	Object	0
Artist Followers	Float64	0.007%
Song ID	Object	0.007%
Genre	Object	0.007%
Release Date	object	0.007%
Weeks Charted	Object	0
Popularity	Float64	0.007%
Danceability	Float64	0.007%
Energy	Float64	0.007%
Loudness	Float64	0.007%
Speechiness	Float64	0.007%
Acousticness	Float64	0.007%
Liveness	Float64	0.007%
Tempo	Float64	0.007%
Duration (ms)	Float64	0.007%
Valence	Float64	0.007%
Chord	Object	0.007%
Joyful	Float64	0.007%
popularity	Float64	0.007%

III. **Data Set Summary Statistics**

Variable Name	Count	Mean	Standard Deviation	Min	25th	50th	75th	Max
Index	1556	779.14	449.59	1	389	780	1169	1556
Highest Charting Position	1545	87.83	58.13	1	37	80	137	200

Number of Times Charted	1545	10.68	16.39	1	1	4	12	142
Streams	1.545000e+03	6.337136e+06	3.375402e+06	4.176083e+06	4.915080e+06	5.269163e+06	6.452492e+06	4.863345e+07
Artist Followers	1.545000e+03	1.471690e+07	1.667579e+07	4.883000e+03	2.123734e+06	6.852509e+06	2.269875e+07	8.333778e+07
Popularity	1545	-8.379766e+06	1.666889e+07	-7.852399e+07	-1.712046e+07	-1.489556e+06	3.526664e+06	4.525569e+07
Danceability	1545	0.696	0.145	0.2	0.6	0.7	0.8	1
Energy	1545	0.64	0.16	0.1	0.5	0.6	0.8	1
Loudness	1545	-6.35	2.51	-25.17	-7.5	-6	-4.71	1.51
Speechiness	1545	0.124	0.11	0.02	0.05	0.08	0.17	0.88
Acousticness	1545	0.249	0.25	0	0.05	0.16	0.39	0.99
Liveness	1545	0.18	0.144	0.02	0.10	0.12	0.22	0.96
Tempo	1545	122.8	29.59	46.72	97.96	122.01	143.86	205.27
Duration (ms)	1545	197940.81	47148.93	30133	169266	193591	218902	588139
Valence	1545	0.52	0.23	0.03	0.34	0.51	0.69	0.98
Joyfulness	1545	9.93	6.70	0.58	5.18	8.70	12.9	70.5

Proportions for Categorical Variables

Category	Frequency	Proportion
Artist (Drake)	29	29/1556 = 1.86%

Category	Frequency	Proportion
Artist (Taylor Swift)	53	53/1556 = 3.40%

Category	Frequency	Proportion
Artist (The Weeknd)	29	29/1556 = 1.86%

Category	Frequency	Proportion
Chord (AB)	130	130/1556 = 8.35%

Category	Frequency	Proportion
Genre ("... rap")	654	654/1556 = 42.03%

Category	Frequency	Proportion
----------	-----------	------------

Genre (R&B)	75	75/1556 = 4.82%
-------------	----	-----------------

Category	Frequency	Proportion
Genre (Latin)	194	194/1556 = 12.46%

Correlation Tables

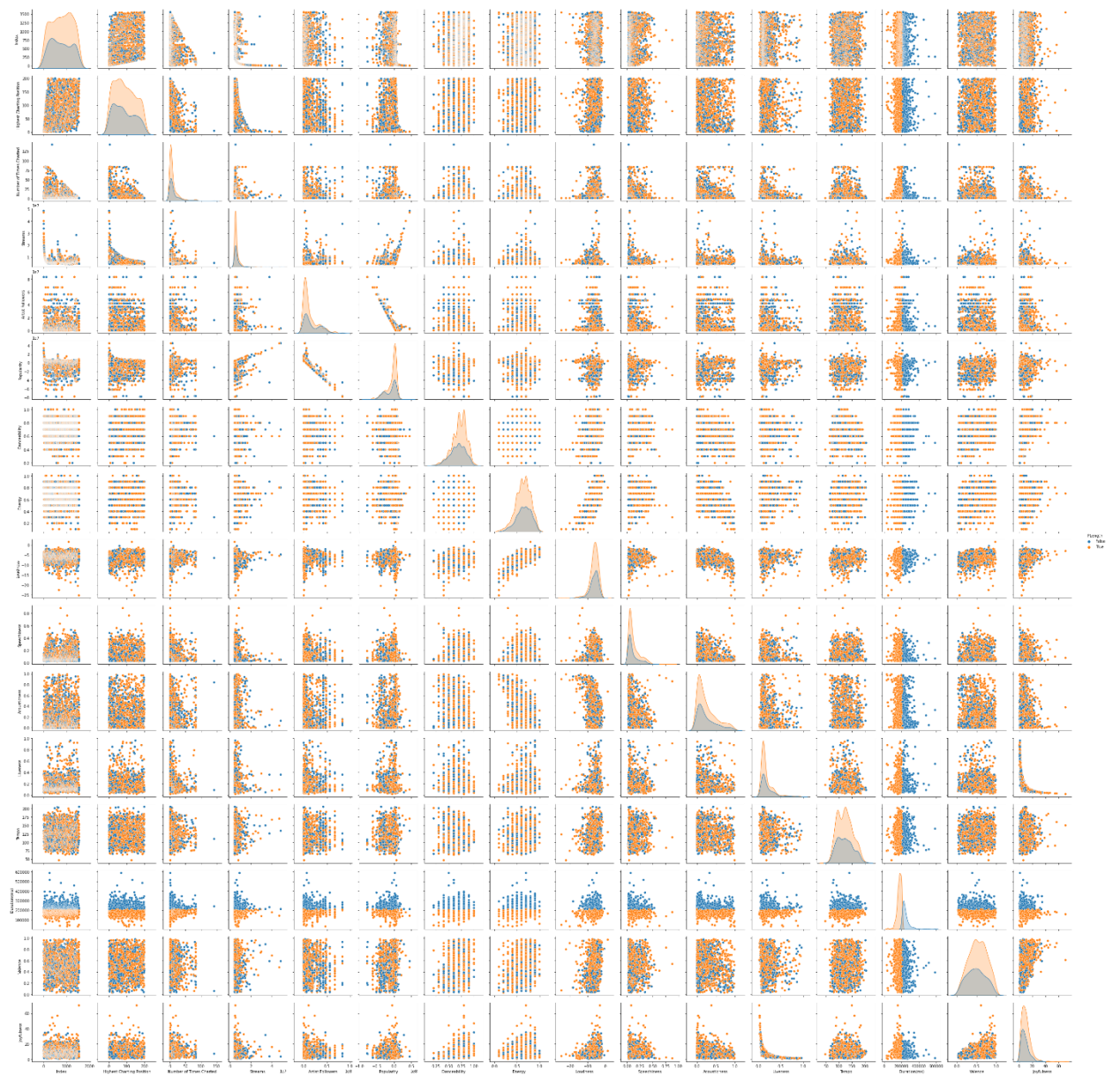
	Ind ex	Hi gh es t Ch ar t in g Po si ti on	Nu m be r of ti m es ch ar t ed	Str ea ms	Art ist Fol lo we rs	Po pul ar it y	Dan cea bilit y	En er gy	Lo ud ne ss	Spe chi nes s	Aco usti cne ss	Liv en es s	Te m po	Du rat io n (m s)	Va le nc e	Joy ful nes s	PL en gt h
Inde x	1.0 00 00 0	0.2 54 34 1	- 0.3 59 69 3	- 0.2 60 03 3	0.0 90 88 7	- 0.1 433 19	0.12 379 2	- 0.0 09 06 5	- 0.0 14 31 4	0.10 923 9	- 0.06 405 5	0.0 29 20 4	0.0 23 62 0	- 0.0 23 48 1	- 0.0 54 86 1	- 0.0 235 96	0.0 34 55 1
Hig hest Cha rtin g Posi tion	0.2 54 34 1	1	- 0.4 17 96 8	- 0.2 95 80 1	- 0.2 33 72 3	0.1 739 94	0.00 865 7	0.0 65 23 6	0.0 32 14 5	0.04 145 9	- 0.01 273 6	0.0 13 61 4	0.0 26 23 2	- 0.0 33 95 6	0.0 45 39 6	- 0.0 004 60	0.0 40 77 0
Nu mbe r of tim es char ted	- 0.3 59 69 3	- 0.4 17 96 8	1	- 0.0 61 63 1	0.0 27 45 8	- 0.0 397 29	0.02 828 8	- 0.0 63 16 3	0.0 31 29 1	- 0.05 999 7	0.04 645 7	- 0.0 58 86 3	- 0.0 48 30 8	0.0 33 98 0	0.0 21 58 8	0.0 482 77	- 0.0 35 93 2
Stre ams	- 0.2 60 03 3	- 0.2 95 80 1	- 0.0 61 63 1	1	0.1 03 25 1	0.0 992 03	- 0.07 882 8	0.0 04 10 5	- 0.0 311 04	- 0.06 261 2	0.03 448 4	0.0 41 39 5	0.0 53 46 6	0.0 15 96 4	0.0 38 19 7	- 0.0 134 33	- 0.0 40 74 5
Arti st	0.0 90	- 0.2 33	0.0 27	0.1 03	1	- 0.9	- 0.09	- 0.0 63	- 0.0 33	- 0.07	0.02 393 6	- 0.0 13	- 0.0 19	0.1 42	- 0.1 09	- 0.0	- 0.1 40

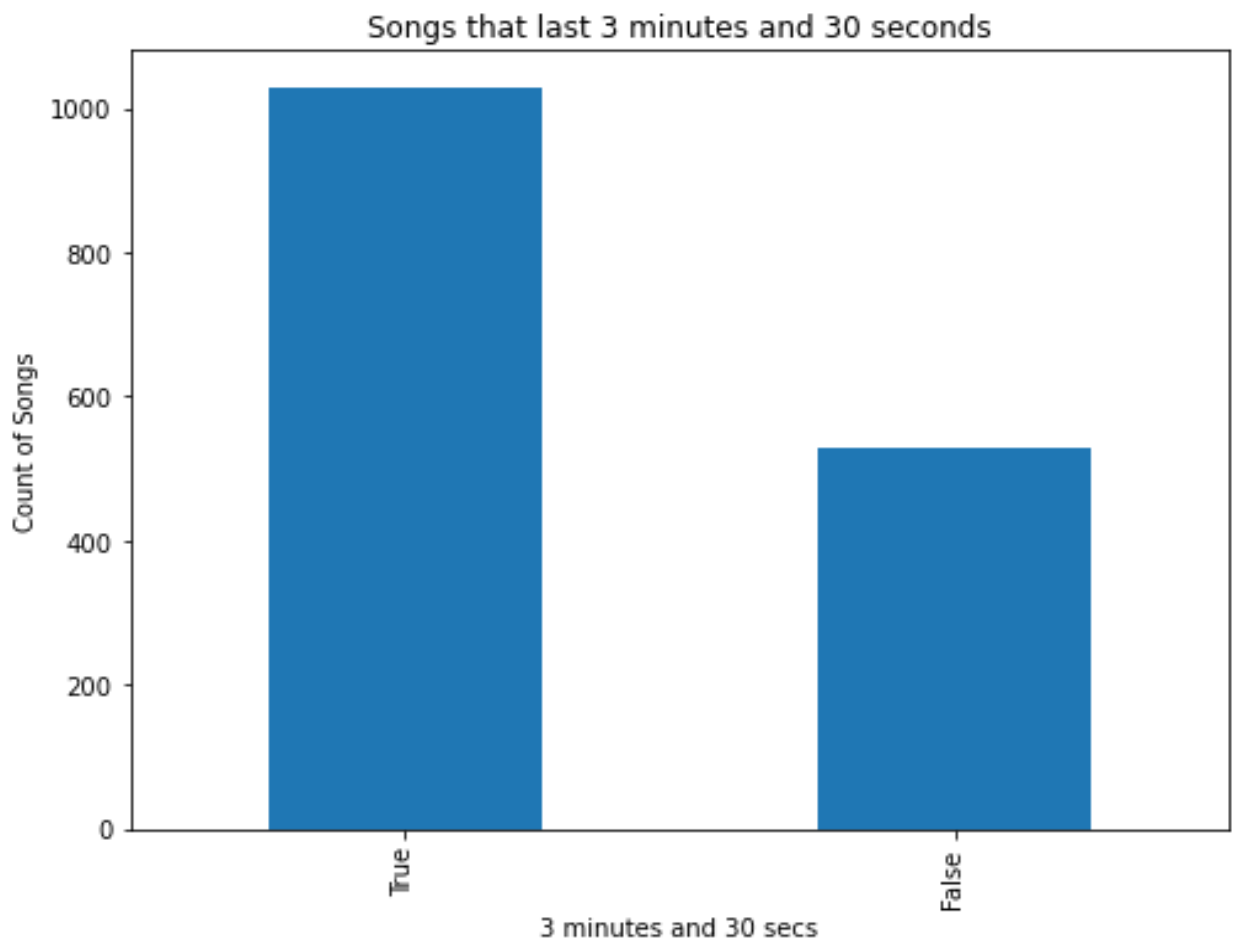
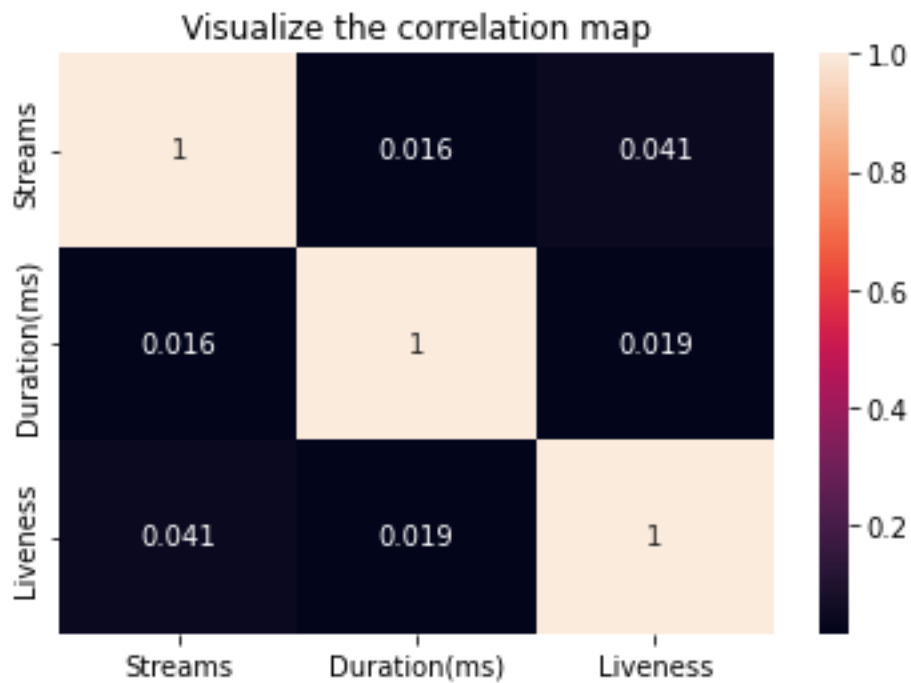
Foll owe rs	88 7	72 3	45 8	25 1		795 06	028 3	80 8	25 3	186 5		75 2	88 0	14 5	03 3	357 03	50 9
Pop ulari ty	- 0.1 43 31 9	0.1 73 99 4	- 0.0 39 72 9	0.0 99 20 3	- 0.9 79 50 6	1	0.07 435 8	0.0 64 66 6	0.0 26 96 9	0.05 921 6	- 0.01 696 3	0.0 22 14 0	0.0 30 71 5	- 0.1 38 97 1	0.1 16 81 3	0.0 329 98	0.1 32 50 5
Dan cea bilit y	0.1 23 79 2	0.0 08 65 7	0.0 28 28 8	- 0.0 78 82 8	- 0.0 90 28 3	0.0 743 58	1	0.1 35 90 0	0.2 33 05 4	0.23 396 7	- 0.30 537 0	- 0.1 14 26 8	- 0.0 97 74 0	0.3 55 50 3	0.3 705 32	0.1 04 70 3	
Ene rgy	- 0.0 09 06 5	0.0 65 23 6	- 0.0 63 16 3	0.0 04 10 5	- 0.0 63 80 8	0.0 646 66	0.13 590 0	1	0.7 22 93 3	0.03 110 7	- 0.52 953 2	0.1 24 48 6	0.1 10 19 2	0.0 56 51 1	0.3 48 25 0	0.0 867 58	- 0.0 44 37 3
Lou dne ss	- 0.0 14 31 4	0.0 32 14 5	0.0 31 29 1	- 0.0 31 10 4	- 0.0 33 25 3	0.0 269 69	0.23 305 4	0.7 22 93 3	1	- 0.01 817 5	- 0.47 673 7	0.0 42 21 2	0.1 04 31 5	0.0 75 26 2	0.2 98 70 8	0.1 385 99	- 0.0 15 16 0
Spe echi ness	0.1 09 23 9	0.0 41 45 9	- 0.0 59 99 7	- 0.0 62 61 2	- 0.0 71 86 5	0.0 592 16	0.23 396 7	0.0 31 10 7	- 0.0 18 17 5	1	- 0.13 168 1	0.0 72 41 5	0.1 11 16 0	- 0.0 90 27 3	0.0 38 59 5	0.0 361 16	0.0 60 25 4
Aco usti cne ss	- 0.0 64 05 5	- 0.0 12 73 6	0.0 46 45 7	0.0 34 48 4	0.0 23 93 6	- 0.0 169 63	- 0.30 537 0	- 0.5 29 53 2	- 0.4 76 73 7	- 0.13 168 1	1	- 0.0 04 60 8	- 0.0 61 47 5	- 0.0 45 88 1	- 0.0 96 06 3	- 0.1 109 08	- 0.0 01 63 2
Live ness	0.0 29 20 4	0.0 13 61 4	- 0.0 58 86 3	0.0 41 39 5	- 0.0 13 75 2	0.0 221 40	- 0.11 426 8	0.1 24 48 6	0.0 42 21 2	0.07 241 5	- 0.00 460 8	1	- 0.0 18 16 2	0.0 18 92 7	0.0 06 84 1	- 0.6 099 12	- 0.0 52 07 2
Tem po	0.0 23 62 0	0.0 26 23 2	- 0.0 48 30 8	0.0 53 46 6	- 0.0 19 88 0	0.0 307 15	- 0.04 220 9	0.1 10 19 2	0.1 04 31 5	0.11 116 0	- 0.06 147 5	- 0.0 18 16 2	1	- 0.0 04 67 1	0.0 57 66 7	0.0 152 18	- 0.0 01 33 6

Dur atio n (ms)	- 0.0 23 48 1	- 0.0 33 95 6	0.0 33 98 0	0.0 15 96 4	0.1 42 14 5	- 0.1 389 71	- 0.09 774 0	0.0 56 51 1	0.0 75 26 2	- 0.09 027 3	- 0.04 588 1	0.0 18 92 7	- 0.0 04 67 1	1	- 0.1 19 45 5	- 0.0 713 83	- 0.7 24 43 0
Vale nce	- 0.0 54 86 1	0.0 45 39 6	0.0 21 58 8	0.0 38 19 7	- 0.1 09 03 3	0.1 168 13	0.35 550 3	0.3 48 25 0	0.2 98 70 8	0.03 859 5	- 0.09 606 3	0.0 06 84 1	0.0 57 66 7	- 0.1 19 45 5	1	0.4 144 21	0.0 95 48 5
Joyf ulne ss	- 0.0 23 59 6	- 0.0 00 46 0	0.0 48 27 7	- 0.0 13 43 3	- 0.0 35 70 3	0.0 329 98	0.37 053 2	0.0 86 75 8	0.1 38 59 9	0.03 611 6	- 0.11 090 8	- 0.6 09 91 2	0.0 15 21 8	- 0.0 71 38 3	0.4 14 42 1	1	0.0 99 81 4
PLe ngt h	0.0 34 55 1	0.0 40 77 0	- 0.0 35 93 2	- 0.0 40 74 5	- 0.1 40 50 9	0.1 325 05	0.10 470 3	- 0.0 44 37 3	- 0.0 15 16 0	0.06 025 4	- 0.00 163 2	- 0.0 52 07 2	- 0.0 01 33 6	- 0.7 24 43 0	0.0 95 48 5	0.0 998 14	1

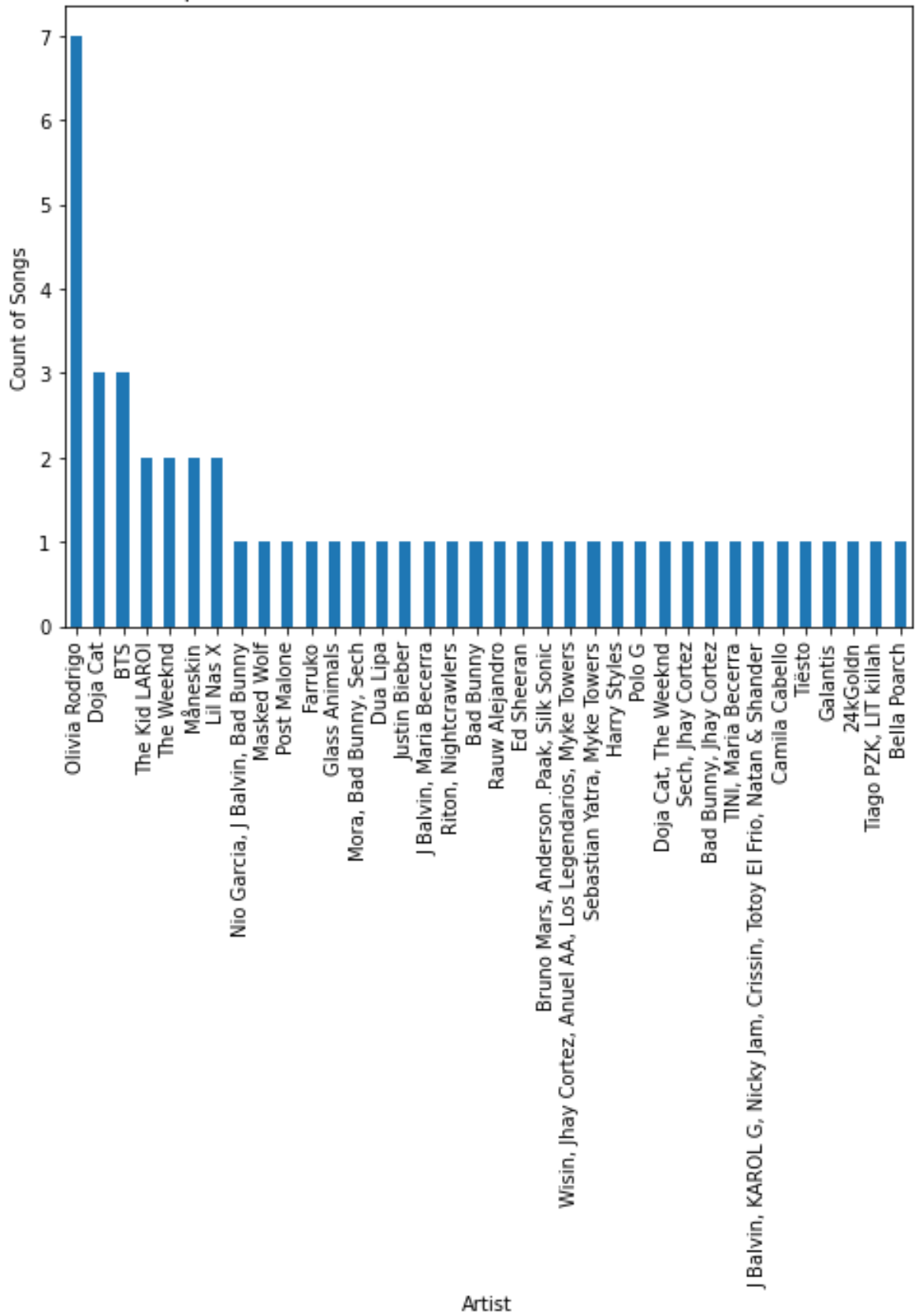
IV. Data Set Graphical Exploration

Pairplot:



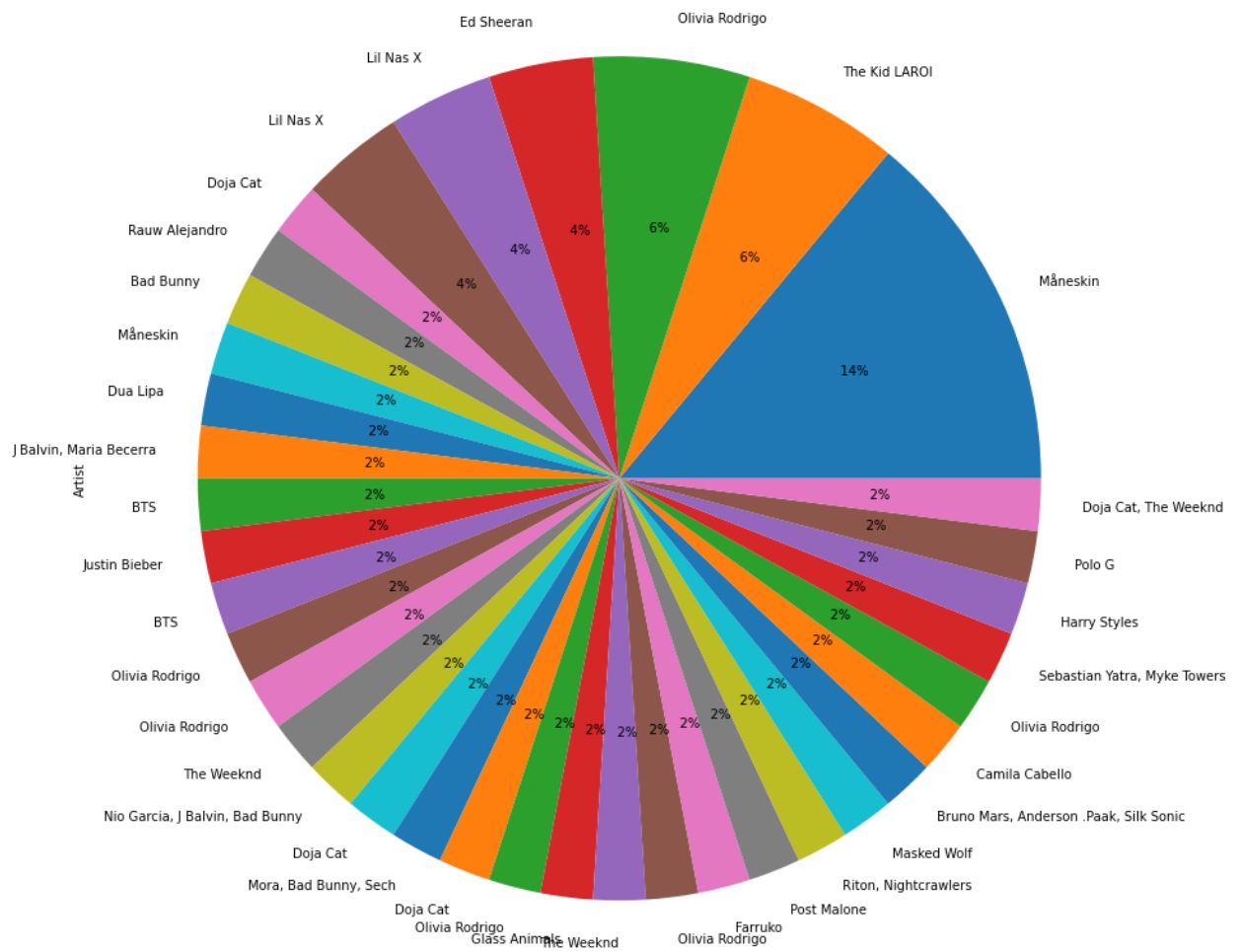


Top 50 Artist whose music last 3 minutes and 30 seconds

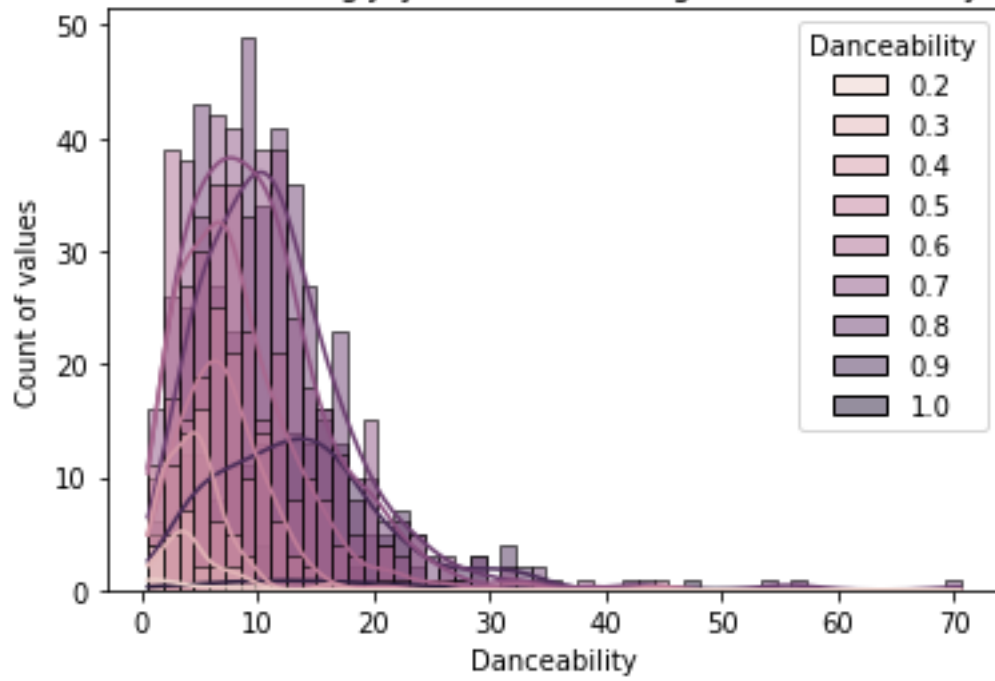


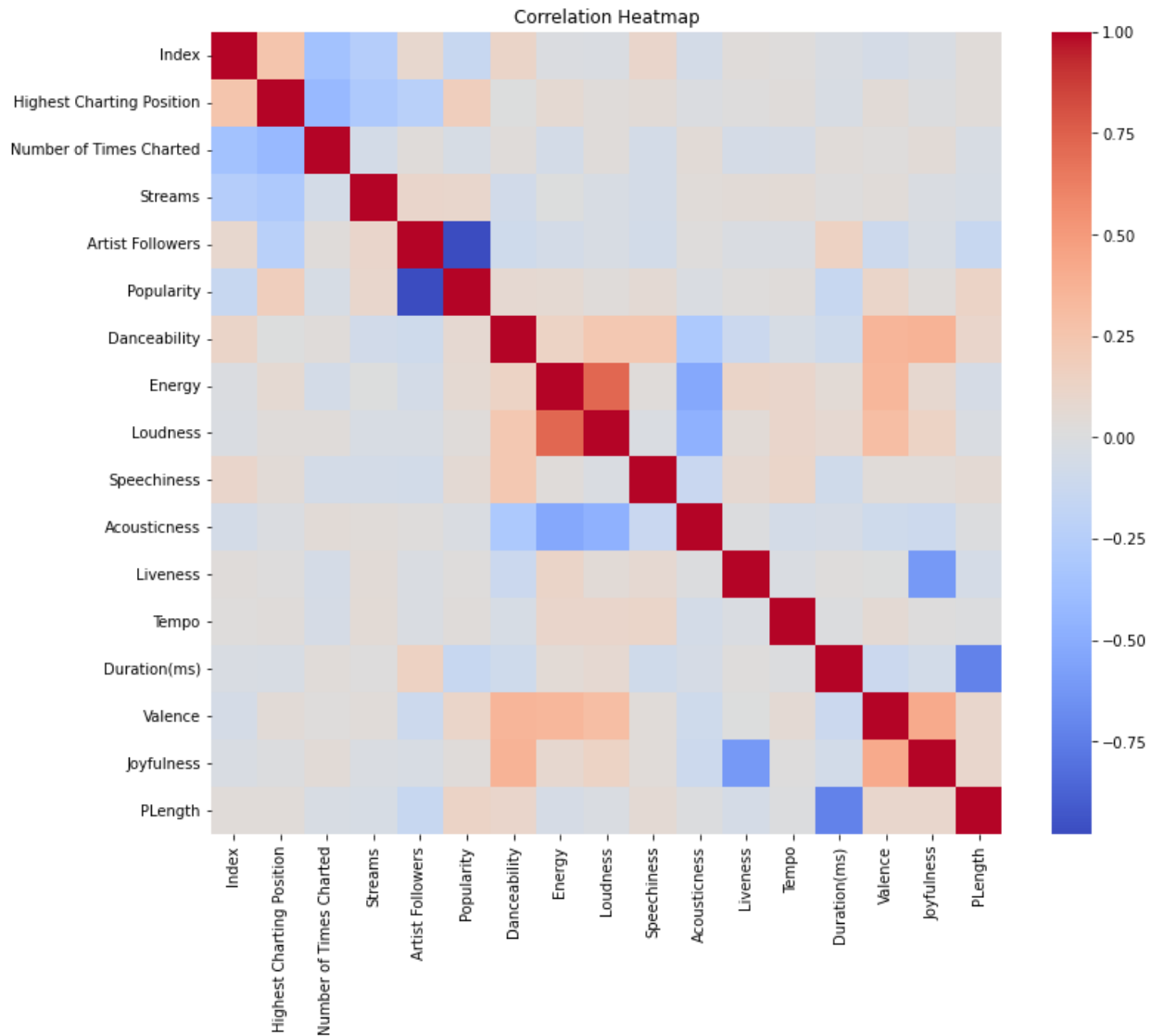
Top 50 Artist v. Length

Artist's music that last 3 minutes and 30 seconds



Understanding Joyfulness of a song with Danceability





- V. Summary:** Through this dataset, we explored the trends of recent pop music and why certain types of songs are popular. Generally, the more danceable a song is, the more likely it will be popular. Also, shorter songs seem to be doing better in this generation. Most of the Top 50 artists have less than 5% of their songs longer than 3 and a half minutes.